GPT4Tools: Teaching Large Language Model to Use Tools via Self-instruction

Year: 2023 | Citations: 276 | Authors: Rui Yang, Lin Song, Yanwei Li

---

Abstract

This paper aims to efficiently enable Large Language Models (LLMs) to use multimodal tools. Advanced proprietary LLMs, such as ChatGPT and GPT-4, have shown great potential for tool usage through sophisticated prompt engineering. Nevertheless, these models typically rely on prohibitive computational costs and publicly inaccessible data. To address these challenges, we propose the GPT4Tools based on self-instruct to enable open-source LLMs, such as LLaMA and OPT, to use tools. It generates an instruction-following dataset by prompting an advanced teacher with various multi-modal contexts. By using the Low-Rank Adaptation (LoRA) optimization, our approach facilitates the open-source LLMs to solve a range of visual problems, including visual comprehension and image generation. Moreover, we provide a benchmark to evaluate the ability of LLMs to use tools, which is performed in both zero-shot and fine-tuning ways. Extensive experiments demonstrate the effectiveness of our method on various language models, which not only significantly improves the accuracy of invoking seen tools, but also enables the zero-shot capacity for unseen tools. The code and demo are available at https://github.com/StevenGrove/GPT4Tools.